

The genome analysis of *Oleiphilus messinensis* ME102 (DSM 13489T) reveals backgrounds of its obligate alkane-devouring marine lifestyle

Toshchakov, Stepan V.; Korzhenkov, Alexei A.; Chernikova, Tatyana; Ferrer, Manuel; Golyshina, Olga; Yakimov, Michail M.; Golyshin, Peter

Marine Genomics

DOI:

[10.1016/j.margen.2017.07.005](https://doi.org/10.1016/j.margen.2017.07.005)

Published: 01/12/2017

Peer reviewed version

[Cyswllt i'r cyhoeddiad / Link to publication](#)

Dyfyniad o'r fersiwn a gyhoeddwyd / Citation for published version (APA):

Toshchakov, S. V., Korzhenkov, A. A., Chernikova, T., Ferrer, M., Golyshina, O., Yakimov, M. M., & Golyshin, P. (2017). The genome analysis of *Oleiphilus messinensis* ME102 (DSM 13489T) reveals backgrounds of its obligate alkane-devouring marine lifestyle. *Marine Genomics*, 36, 41-47. <https://doi.org/10.1016/j.margen.2017.07.005>

Hawliau Cyffredinol / General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

The genome analysis of *Oleiphilus messinensis* ME102 (DSM 13489^T) reveals backgrounds of its obligate alkane-devouring marine lifestyle

Stepan V. Toshchakov¹, Alexei A. Korzhenkov¹, Tatyana N. Chernikova², Manuel Ferrer³, Olga V. Golyshina², Michail M. Yakimov^{1,4} and Peter N. Golyshin²

¹ Immanuel Kant Baltic Federal University, 236040 Kaliningrad, Russia

² School of Biological Sciences, Bangor University, LL57 2UW Bangor, Gwynedd, UK

³ Institute of Catalysis CSIC, Campus Cantoblanco, 28049 Madrid, Spain

⁴ Institute for Coastal Marine Environment, CNR, 98122 Messina, Italy

Keywords:

Oceanospirillales*, marine alkane-degrading bacteria, hydrocarbonoclastic, *Oleiphilus messinensis

Abstract

Marine bacterium *Oleiphilus messinensis* ME102 (DSM 13489^T) isolated from the sediments of the harbor of Messina (Italy) is a member of the order *Oceanospirillales*, class *Gammaproteobacteria*, representing the physiological group of marine obligate hydrocarbonoclastic bacteria (OHCB) alongside the members of the genera *Alcanivorax*, *Oleispira*, *Thalassolituus*, *Cycloclasticus* and *Neptunomonas*. These organisms play a crucial role in the natural environmental cleanup in marine systems. Despite having the largest genome (6.379.281 bp) among OHCB, *O. messinensis* exhibits a very narrow substrate profile. The alkane metabolism is pre-determined by three loci encoding for two P450 family monooxygenases, one of which formed a cassette with ferredoxin and alcohol dehydrogenase encoding genes and alkane monooxygenase (AlkB) gene clustered with two genes for rubredoxins and NAD⁺-dependent rubredoxin reductase. Its genome contains the largest numbers of genomic islands (15) and mobile genetic elements (140), as compared with more streamlined genomes of its OHCB counterparts. Among hydrocarbon-degrading *Oceanospirillales*, *O. messinensis* encodes the largest array of proteins involved in the signal transduction for sensing, and responding to the, environmental stimuli (345 vs 170 in *Oleispira antarctica*, the bacterium with the second highest number). This must be an important trait to adapt to the conditions in marine sediments with a high physico-chemical patchiness and heterogeneity as compared to those in the water column.

1. Introduction

Oleiphilus messinensis is one of the organisms that represent the functional group of so-called obligate marine hydrocarbonoclastic bacteria (OHCB), which plays a pivotal role in the degradation of petroleum constituents in the sea (Yakimov et al., 2007). The strain *O. messinensis* ME102 (DSM 13489^T) was isolated from the sediments of the Messina harbor, which is severely affected by the ferry traffic, after the enrichment of the sample with tetradecane in the artificial seawater medium ONR7a. *O. messinensis* was described to represent a new species within the new genus *Oleiphilus*, within the novel family *Oleiphilacea* of *Oceanosprillales* (*Gammaproteobacteria*) (Golyshin et al., 2002). These Gram-negative aerobic bacteria have a very restricted substrate range, consistent with their OHCB designation, preferring aliphatic hydrocarbons, fatty acids and alcohols, as carbon and energy sources over sugars and amino acids (Golyshin et al., 2002). Figure 1 depicts the placement of *O. messinensis* pointing at its relatively distant placement on the phylogenetic tree with other *Oceanospirillales*. Worth attention, the genus *Oleiphilus* is currently only represented by a single isolate (type strain) with another one to share 98% SSU rRNA sequence identity (GenBank Acc Nr FJ845394), pointing at a rather endemic nature of this particular species. This is in the stark contrast with other OHCBs that include very ubiquitous *Alcanivorax* spp. (Yakimov et al., 1998, Schneiker et al., 2006), *Oleispira* spp. important in polar and deep (cold) marine environments (Yakimov et al., 2003; Kube et al., 2013), *Thalassolituus* spp., an oil-degrader occupying various marine niches, including estuarine waters (Yakimov et al., 2004; McKew et al., 2007; Golyshin et al., 2013) and PAH-degrading specialists from the genus *Cycloclasticus* (Dyksterhouse et al., 1995; Geiselbrecht et al., 1998; Lai et al., 2012; Messina et al., 2016). Here, we report on the genome-based analysis of obligate marine hydrocarbonoclastic bacterium, *Oleiphilus messinensis* ME102. Genome analysis revealed exceptional genome plasticity of

ME102, showing an unprecedented abundance of mobile elements for a member of the *Oceanospirillales*, which could potentially play an important role in the of genome regulatory circuits.

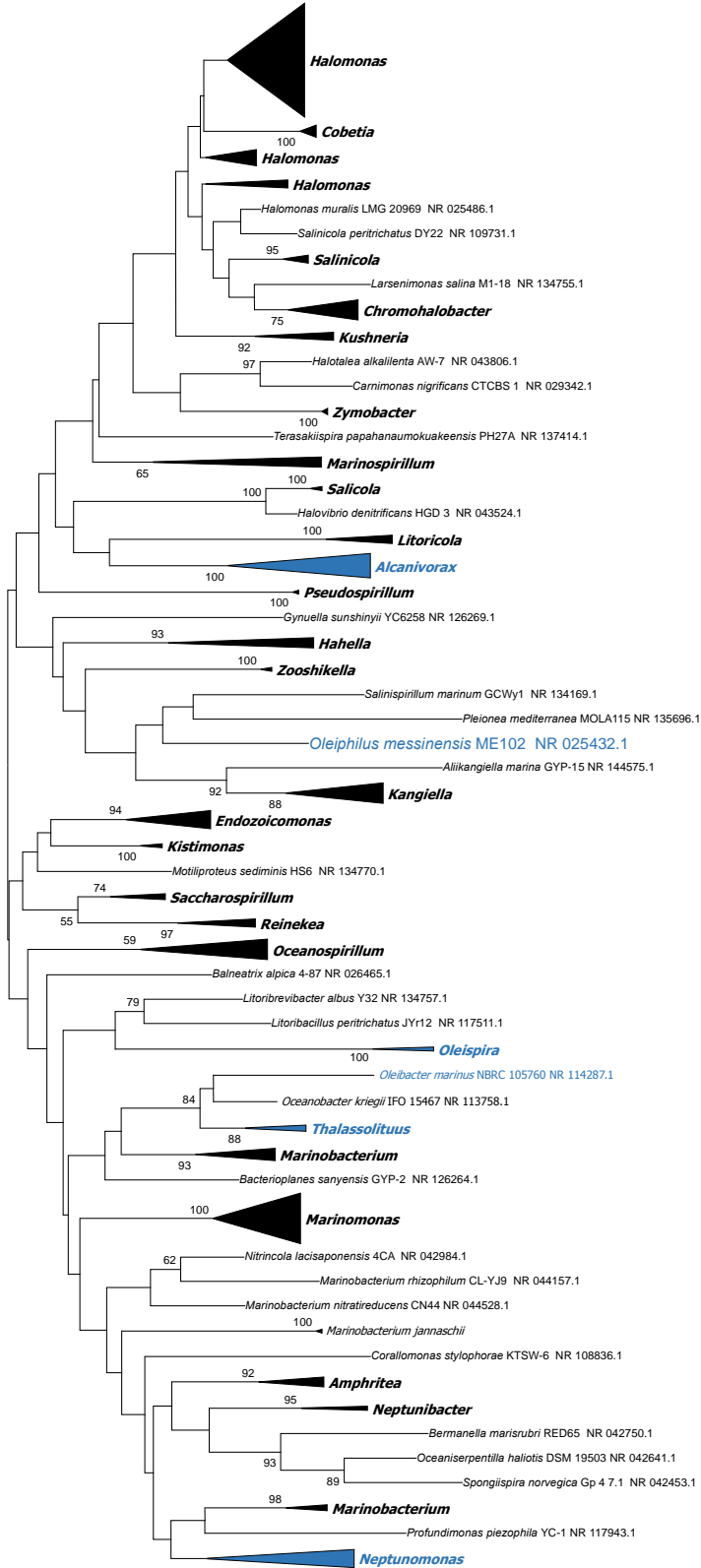


Fig. 1. Phylogenetic position of *O. messinensis* ME102 and other OHCB (marked in blue) among 250 species of *Oceanospirillales*, as revealed by SSU rRNA gene sequence analysis. The evolutionary relationships were inferred using the Neighbor-Joining method (Saitou and Nei, 1987). The optimal tree with the sum of branch length = 3.41994391 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test > 50% (1000 replicates) are shown next to the branches (Felsenstein, 1985). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Jukes-Cantor method (Jukes and Cantor, 1969) and are in the units of the number of base substitutions per site. The rate variation among sites was modeled with a gamma distribution (shape parameter = 1). The analysis involved 251 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 464 positions in the final dataset. Evolutionary analyses were conducted in MEGA6 (Tamura et al., 2013).

2.1 General features of *O. messinensis* genome.

The genome of *O. messinensis* was sequenced using hybrid approach using Roche 454 and Illumina sequencing technologies. Assembly was performed with Newbler and Phred/Phrap/Consed *de novo* assemblers resulting in 6.38 Mb circular chromosome sequenced with 119x read coverage. Genome size of *O. messinensis* is the largest among other hydrocarbonoclastic *Oceanospirillales* reported earlier (Figure 2). Average GC content was 47.8 %. Genome contains 5502 protein-coding genes, 4081 (74,1 %) of which were assigned with function. Five complete ribosomal operons have identical structure: 16S – tRNA-Ile – tRNA-Ala – 23S – 5S. *O. messinensis* genome was significantly enriched with mobile elements, more than 50% of which were active (Table 1). Analysis of bacteriophage-related genes with PHASTER (Arndt et al., 2016) revealed two putative prophage regions of 21.5 kb total.

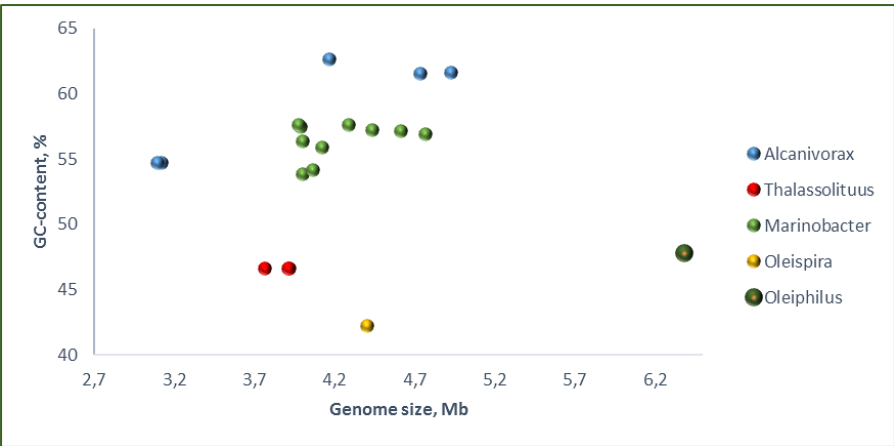


Fig. 2 Genome size vs. GC-content in complete genomes of gammaproteobacterial genera with hydrocarbonoclastic representatives.

Table 1. Genome and environmental features of *Oleiphilus messinensis*

Feature	Description
Current classification	Domain <i>Bacteria</i> Phylum <i>Proteobacteria</i> Class <i>Gammaproteobacteria</i> Order <i>Oceanospirillales</i> Family <i>Oleiphilaceae</i> Genus <i>Oleiphilus</i> Species <i>Oleiphilus messinensis</i> Type strain ME102 ^T
Biosample ref	SAMN06234459 (NCBI)
Collection date	21-Jun-1998
Geographical location name	Italy, Messina harbor, 38.19N 15.56 E
Geographical coordinates	38.19 N 15.56 E
env_biome	sea
env_feature	harbor
env_material	water/sediment
depth, m	8
Biotic relationship	free living
Relation to oxygen	aerobe
Sequencing project ref	PRJNA362330 (NCBI)
Sequencing method	Roche 454, Illumina GAx
Sequencing center	Fidelity Systems
Assembly method	Newbler v. 2.6; Phred/Phrap/Consed v. 23.0
Coverage	119x
Number of replicons	1
Finishing level	Finished
Genome size, bp	6379281
GC content, %	47.8
Genes	5502
Pseudogenes	53
RNA genes	74
rRNA	15 (5 operons)
tRNA	59
GI number	15
GI length (% share)	218392 (3.4)
Mobile elements	142
complete ORFs	85
partial ORFs	57
Number of different IS families	14
Prophage regions	2
intact	-
partial	2
total length (kb)	21.5

2.2 Genome-based phylogenetic position

Phylogenetic reconstruction of marine hydrocarbonoclastic Gammaproteobacteria using 32 ribosomal proteins showed that *O. messinensis* represents a deep lineage of *Oceanospirillales*. *Hahellaceae* family with the only sequenced genome of *Hahella chejuensis* KCTC 2396 (Jeong et al., 2005) was the closest relative to *Oleiphilus* spp. While 16S rRNA identity of ME102 and KCTC2396 was on the border of family threshold (92 %), these microorganisms are characterized by quite different substrate preferences and other phenotypic characteristics (Lee et al., 2001; Golyshin et al., 2002). Genomic data also supports their classification as type strains of different families. In addition to significant GC difference (6.07%), mean two-way amino acid identity of two *in silico* proteomes was 53.17 % (Supplementary Figure 1) which is typical for inter-family comparisons (Konstantinidis & Tiedje, 2005).

Interestingly, the group of *Marinobacter* species, currently classified within the family *Alteromonadaceae* of *Alteromonadales*, formed monophyletic group, which was more closely related to *Oleiphilaceae*, than other hydrocarbonoclastic *Oceanospirillales*.

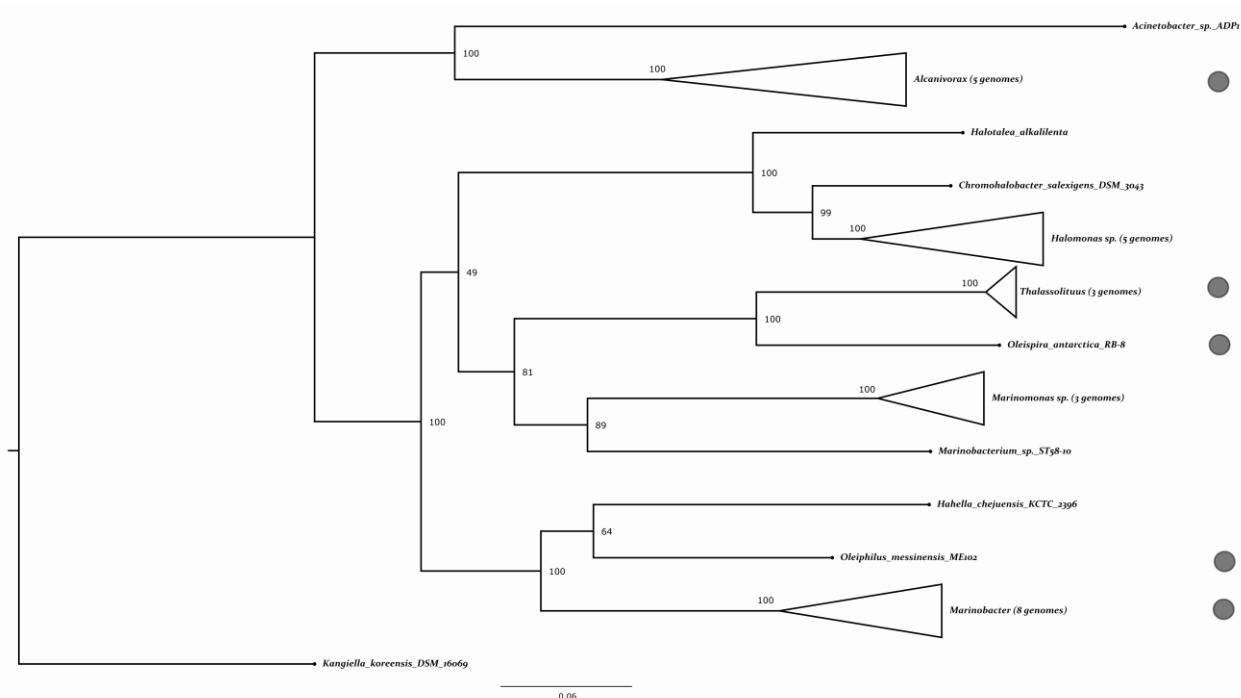


Figure 3. Maximum-likelihood phylogenetic tree of representatives of *Oceanospirillales* family based on concatenated alignment of 22 ribosomal proteins. For phylogenetic analysis, we took complete genomes of free-living gammaproteobacteria from *Oceanospirillales*, *Alteromonadales* (30 and 98 respectively) and *Acinetobacter* sp. ADP1 (order *Pseudomonadales*). Sequences of 22 ribosomal proteins for each genome were accessed using *hmmsearch* script from HMMER 3.1b2 package (e-value < 10⁻⁷, query coverage > 70%) (Finn et al., 2015). 111 genomes possessing all queried ribosomal proteins were used in the further analysis. Sequences for each protein were aligned with Clustal Omega (ver. 1.2.1) and concatenated to one alignment (Sievers and Higgins, 2014). Sites having more than 95% gaps were deleted. Phylogenetic tree was constructed using maximum likelihood method in RAXML (version 8.2.4) with following options: protein model - LG (best scoring), rate heterogeneity model - GAMMA, 100 bootstraps (Stamatakis, 2014). On the basis of initial tree analysis (Supplementary Figure 2) *Kangiella* sp. was chosen as an outgroup. For the final tree construction 31 sequences of OHCB were realigned with Clustal Omega. Tree was constructed in RAXML using parameters described above. Lineages including OHCB marked with grey circles.

2.3 Alkane metabolism

As a typical representative of marine obligate hydrocarbonoclastic bacteria (OHCB) group, *O. messinensis* ME102^T grows preferentially on aliphatic hydrocarbons, alkanolates and alkanols, as sole carbon and energy sources (Golyshin et al., 2002). Expectedly, its genome encodes many genes and operons presumably involved in hydrocarbon metabolism.

The first step of alkane utilization involves terminal hydroxylation of alkanes by alkane-1-monooxygenase. OLMES_3728 protein, orthologous to AlkB₁ alkane hydroxylase of *A. borkumensis* SK2 (Schneiker et al., 2006) clusters with two rubredoxin genes (OLMES_3726

and OLMES_3727), NAD⁺-dependent rubredoxin reductase (OLMES_3725) and transcriptional regulator OLMES_3724 (Figure 4a). OLMES_3724 belongs to LuxR family of transcriptional regulators and has a significant level of homology with AlkS regulator, known to activate transcription of *alk* operon in the presence of alkanes (Kok et al., 1989; Canosa et al., 2000). Thus, OLMES_3724 – OLMES_3728 *alkBFGTS* cluster forms a complete minimal set of genes necessary for primary oxidation of medium-chain alkanes. Another alkane-1-monooxygenase gene, presented in closely related to ME102^T OHCB *A. borkumensis* SK2, *alkB*₂ was apparently lost during evolution.

A. borkumensis SK2 as well as other OHCB also involves P450-like cytochromes for the oxidation of medium-chain alkanes (Sabirova et al., 2007). *O. messinensis* ME102 harbors 3 loci, containing proteins, orthologous to P450 of *A. borkumensis*, known to be upregulated in the presence of isoprenoid hydrocarbons (Schneiker et al., 2006). P450(c)-like protein OLMES_4335 clusters with AraC transcriptional regulator (OLMES_4332), ferredoxin (OLMES_4334) and OLMES_4336 alcohol dehydrogenase, similar to AlkJ alcohol (alkanol) dehydrogenase. P450(b)-like protein seems to be disrupted by the insertion of IS4 family IS element, which, considering the completeness of transposase gene and inverted repeats happened recently (Figure 4c). OLMES_5166 protein, orthologous to P450(a) was found to be isolated from other genes involved in alkane metabolism, but positioned next to OLMES_5167 cupredoxin. Decreased diversity of *alkB* genes and P450 cytochromes can explain narrower spectra of hydrocarbons, which can be utilized by *O. messinensis*, as compared to *A. borkumensis* (Golyshin et al., 2002; Yakimov et al., 1998).

Hydroxylation of long-chain alkanes (18 carbon atoms and longer) requires other alkane hydroxylation systems, involving Alma (Throne-Holst et al., 2007) or LadA (Wentzel et al., 2007) hydroxylases. ME102^T is able to utilize alkanes up to 20 carbon atoms long (Golyshin

et al., 2002). That phenotype may be explained by the fact that ME102^T contains FAD-binding monooxygenase OLMES_5009, having high level of homology (54% amino acid identity) with AlmA long-chain alkane monooxygenase of *Acinetobacter*.

Further steps in alkane catabolism involve successive alkanol oxidation steps resulting in acyl-CoA, which then metabolized through beta-oxidation. In experimentally characterized *Pseudomonas putida* *alk* operon it involves AlkJ alkanol dehydrogenase, AlkH aldehyde dehydrogenase and AlkK fatty-acid CoA ligase (van Beilen et al., 2001). In *O. messinensis* these genes were not clustered around central *alkBFGTS* operon, but rather scattered across the genome. Thus, putative AlkJ alcohol (alkanol) dehydrogenase OLMES_4336 is located near P450 cytochrome. ME102 possess several NAD-dependent aldehyde dehydrogenases, which can be involved in alkane-derived aldehyde oxidation. OLMES_0331 medium-chain-fatty-acid CoA ligase can act as AlkK protein, providing acyl-CoA for further beta-oxidation. As it was previously reported, majority of above gene clusters in *O. messinensis* are situated on horizontally acquired “catabolic transposons”, which is a very common phenomenon in marine oil-degrading bacteria (Yakimov et al., 2007).

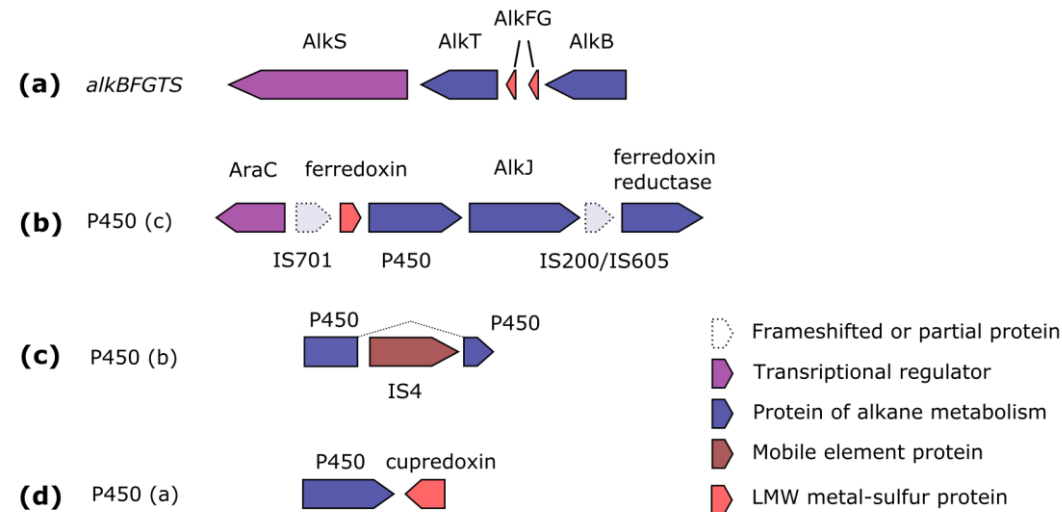


Figure 4. Schematic representation of gene clusters involved in alkane degradation in *O. messinensis*. See text and legend for details.

Genome analysis also revealed coding potential of *O. messinensis* for the utilization of other hydrocarbon derivatives. OLMES_3894 haloalkane-dehalogenase-like protein, which does not have orthologs in other OHCB can possibly convert 1-haloalkanes to primary alcohol. Despite its activity requires further experimental validation, it can be inferred that this enzyme might determine a biotechnological potential of ME102^T in bioremediation of toxic environmental pollutants.

2.4 Genomic islands and mobile genetic elements

Genome of *O. messinensis* contains 15 genomic islands (GIs), most of which are associated with active mobile-element related genes (Figure 5). Total length of all predicted GI was 218392 bp, equivalent to 3.4% of the genome. Analysis of genes, associated with genomic islands, allowed to find out that most of fixed HGT events are of adaptive nature. Most of identified GIs contains genes involved in adaptation to environmental stimuli in marine environments. 9 of 15 GIs contain genetic determinants for bacterial defense and competition systems, such as RelB/RelE (OLMES_0561-OLMES_0562; OLMES_4307-OLMES_4308), VapB/VapC (OLMES_3971-OLMES_3972; OLMES_4311-OLMES_4312), MazE/MazF (OLMES_5329-OLMES_5330), Phd-Doc (OLMES_3985-OLMES_3986) toxin-antitoxin systems; type II (OLMES_5494-OLMES_5495) and type III (OLMES_3674-OLMES_3675) restriction-modification systems. GI4 contains at least two Rhs family proteins, known to be involved in inter- and intraspecies competition (Murdoch et al., 2017).

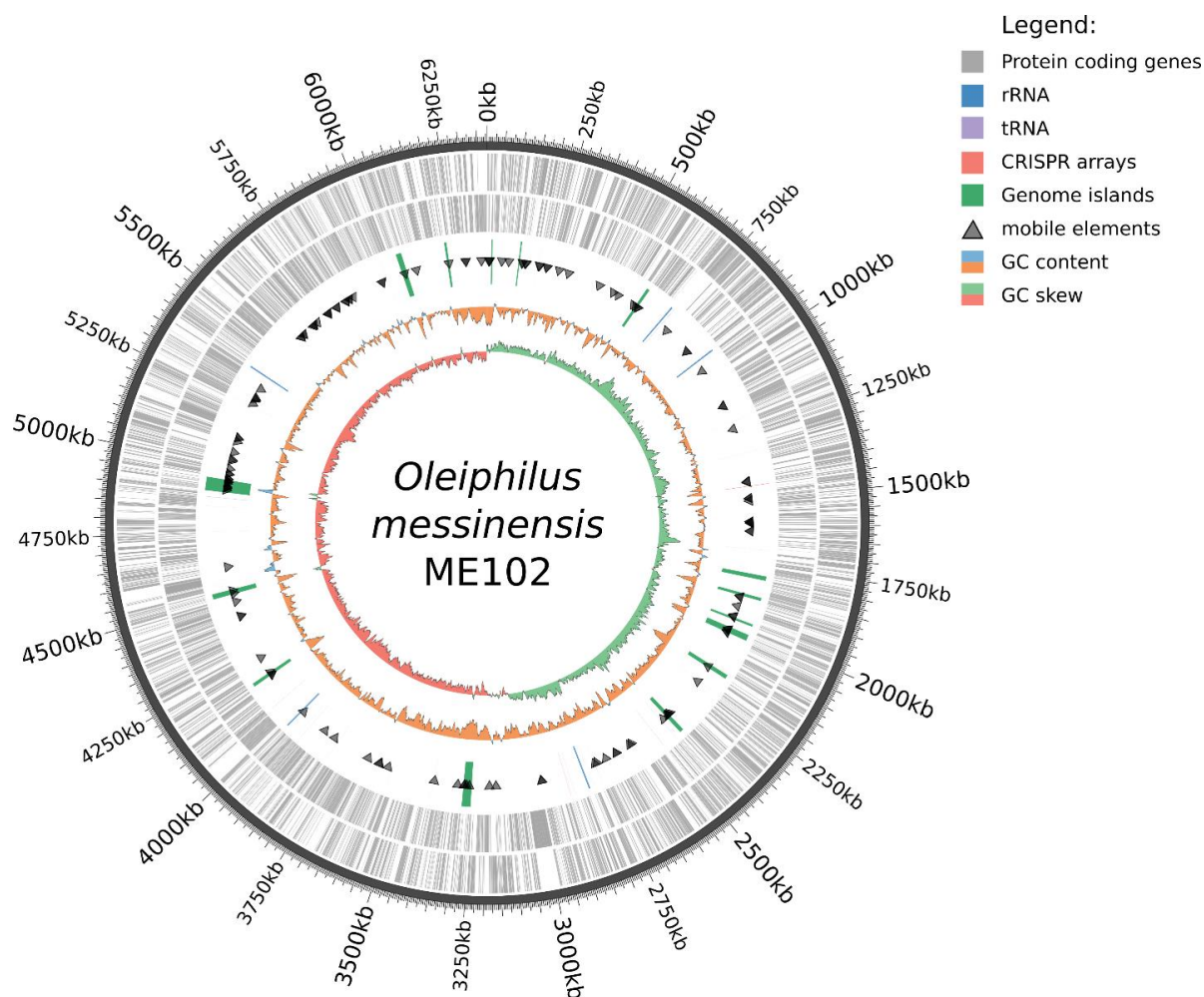


Figure 5. The genome map of *O. messinensis* ME102.

The large island (GI10) spanning 30.5 kbp contains several genes of widespread bacterial chemotaxis operon (OLMES_2869-OLMES_2876) including CheA histidine kinase, CheY and CheW response regulators and methyl-accepting chemotaxis protein MCP (Wuichet et al., 2007) as well as some genes required for mercury fitness, OLMES_2880 – OLMES_2882 (Sone et al., 2013).

The biggest GI identified in *O. messinensis* genome, the GI13, spanning over 31.7 kbp, possesses genes that might be essential for the phenotype of the organism. This GI contains cytochrome P450 alkane hydroxylase, OLMES_4335 involved in terminal hydroxylation of alkane hydrocarbons, located next to alcohol dehydrogenase (OLMES_4336), which might play a role further steps of alkane utilization.

O. messinensis genome harbours 142 functional and inactivated mobile elements, occupying totally over 192 kbp. Analysis of transposase diversity performed with ISSaga (Varani et al., 2011) showed that they belong to at least 14 different families (Supplementary Table 1) which is rather exceptional among other *Oceanospirillales*, typically possessing 1-4 different IS families. Search for triggers of such genome plasticity among *Oceanospirillales* didn't reveal any significant result, therefore we suggest that expansion of mobile genetic elements might be beneficial to adaptation of changeable marine environments not only by direct transfer of metabolically important genes, but also by attenuation of transcriptional activity of adjacent genes by action of outward-oriented promoters, carried by certain type of IS elements (Vandecraen et al., 2017). IS1380-related elements which seemed to be most active in ME102 genome represented by 18 complete transposase ORFs, known to modulate the expression of genes involved in metronidazole (Soki et al., 2006), carbapenem (Kato et al., 2003) and blasticidine resistance (Lartigue et al., 2006). IS3 elements, also widespread and active in ME102 (17 functional copies), are shown to activate genes involved in acetate (Treves et al., 1998) or citrate utilization (Blount et al., 2012).

2.5 *Oleiphilus messinensis* has a distinguished standing among oil-degrading *Oceanospirillales*

We performed the analysis of distribution of Clusters of Orthologous Groups of proteins (COGs) in fully assembled genomes of oil-degrading *Oceanospirillales*. Interestingly, in *O. messinensis*, only 67.5 % of all proteins were attributed to COGs, whereas in other marine alkane-degraders this value was in the range of 72-83 % with *Alcanivorax*, the organism with the most streamlined genome having >83% proteins in COGs. Unsurprisingly therefore that

Oleiphilus had the largest percentage of proteins in the category R (general function prediction only) (Figure 6). Likely, due to its largest genome size among analyzed genomes, the percentage of proteins of categories C and E (energy production and conversion and amino acid transport and metabolism, correspondingly) were underrepresented. As discussed above, the absolute number of mobile genetic elements (functional category X) in *Oleiphilus* is remarkably high, with 142, by far outperforming next-following *Oleispira antarctica* with just 53 active or inactivated transposases in the genome of the latter. This may be a result of *Oleiphilus* spp. being native inhabitants of the marine sediment, where they inevitably comes into a physical contact with other microorganisms, which is essential for the DNA acquisition/lateral gene transfer.

The census of proteins important in the signal transduction (in particular, with the domains HisK, MCP, GGDEF, GGDEF+EAL, EAL, HD-GYP, AC1, AC3, STYK, ABC1, PP2C and RRs) has revealed *O. messinensis* having by far the largest number of these proteins (345) with *Oleispira antarctica* (the runner-up) having just 170. This is also reflected in the Figure 6, which shows the signal-transduction-related proteins having a highest percentage in total *in silico* proteomes of alkane-degrading *Oceanospirillales* (combined categories T+TK+KT+NT). This suggests that *Oleiphilus* may have a very sophisticated means for sensing the environment and react to the changes therein. This again points at the origin of *Oleiphilus* spp. from marine sediments, where the spatial environmental (physico-chemical) patchiness and heterogeneity are much more pronounced than in a relatively homogenous and constant water column.

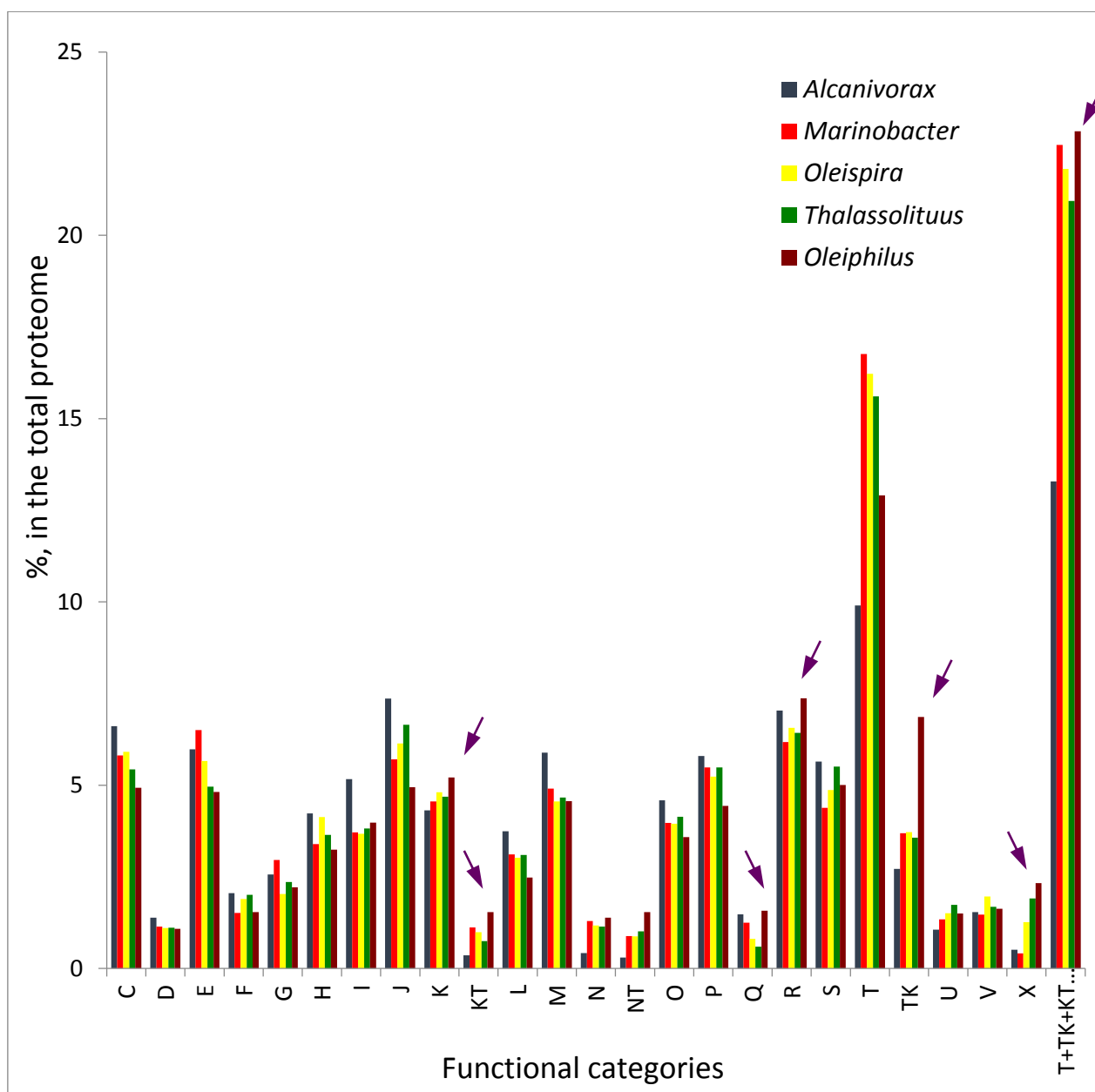


Fig. 6. Distribution of proteins per functional categories of Clusters of Orthologous Groups of proteins (COGs) in genomes of the most significant marine oil degraders. The functional classification of the COGs is shown following functional categories: C, Energy production and conversion; D, Cell cycle control, cell division, chromosome partitioning; E, Amino acid transport and metabolism; F, Nucleotide transport and metabolism; G, Carbohydrate transport and metabolism; H, Coenzyme transport and metabolism; I, Lipid transport and metabolism; J, Translation, ribosomal structure and biogenesis; K, Transcription; L, Replication, recombination and repair; M, Cell wall/membrane biogenesis; N, Cell motility; O, Posttranslational modification, protein

turnover, chaperones; P, Inorganic ion transport and metabolism; Q, Secondary metabolites biosynthesis, transport and catabolism; R, General function prediction only; S, Function unknown; T, Signal transduction mechanisms; U, Intracellular trafficking and secretion; V, Defense mechanisms (Galperin et al., 2015). Purple arrowheads indicate functional categories of COGs overrepresented in *Oleiphilus* as compared with other hydrocarbon-degrading *Oceanospirillales*.

CONCLUSION

The recent blowout of Deepwater Horizon platform in the Gulf of Mexico was followed by one of the largest offshore oil spills with ~4 million barrels crude oil released into the sea.

Countless studies have reported a number of autochthonous marine hydrocarbon-degrading bacteria from the order *Oceanospirillales* have been identified to be actively involved in the oil degradation. The bloom of these organisms resulted in the rapid degradation of many oil constituents, hence highlighting the importance of OHCB in bioremediation of marine environment and a necessity of comprehensive studies to unveil the genomic and physiological backgrounds of hydrocarbonoclastic lifestyle of OHCB.

Despite having the largest genome among OHCB, *O. messinensis* exhibits a very narrow substrate profile and contains the largest numbers of mobile genetic elements, as compared with more streamlined genomes of other OHCB counterparts. With this study, we extended the list of OHCB whose genomes were fully sequenced to further expand our understanding of the efficiency and functional redundancy in hydrocarbon utilization by OHCB, the metabolic routes underlying their special hydrocarbon diet and their ecological success.

CONFLICTS OF INTERESTS

Authors declare no conflicts of interests

300 **ACKNOWLEDGMENTS**

301 This project has received funding from the European Union's Horizon 2020 research and
 302 innovation program [Blue Growth: Unlocking the potential of Seas and Oceans] under grant
 303 agreement No [634486] (Project INMARE). This research was also supported by the European
 304 Community FP7 Projects MAMBA (FP7-KBBE-2008-226977) and KILLSPILL (FP7-KBBE-2012-
 305 312139). The present investigation was also funded by the ERA NET IB2 Project MetaCat,
 306 grant number ERA-IB-14-030, through UK Biotechnology and Biological Sciences Research
 307 Council (BBSRC), grant BB/M029085/1.

308

309

310 **REFERENCES**

311

- 312 Arndt, D., Grant, J.R., Marcu, A., Sajed, T., Pon, A., Liang, Y., Wishart, D.S., 2016. PHASTER: a better, faster version
 313 of the PHAST phage search tool. *Nucleic Acids Res.* 44, W16-21. doi:10.1093/nar/gkw387
- 314 Blount, Z.D., Barrick, J.E., Davidson, C.J., Lenski, R.E., 2012. Genomic analysis of a key innovation in an
 315 experimental *Escherichia coli* population. *Nature* 489, 513–518. doi:10.1038/nature11514
- 316 Canosa, I., Sánchez-Romero, J.M., Yuste, L., Rojo, F., 2000. A positive feedback mechanism controls expression of
 317 AlkS, the transcriptional regulator of the *Pseudomonas oleovorans* alkane degradation pathway. *Mol.*
 318 *Microbiol.* 35, 791–9.
- 319 Dyksterhouse, S.E., Gray, J.P., Herwig, R.P., Lara, J.C., Staley, J.T., 1995. *Cycloclasticus pugetii* gen. nov., sp. nov.,
 320 an Aromatic Hydrocarbon-Degrading Bacterium from Marine Sediments. *Int. J. Syst. Bacteriol.* 45, 116–123.
 321 doi:10.1099/00207713-45-1-116

322 Felsenstein, J., 1985. Confidence Limits on Phylogenies: An Approach Using the Bootstrap. *Evolution* (N. Y). 39,
323 783. doi:10.2307/2408678

324 Finn, R.D., Clements, J., Arndt, W., Miller, B.L., Wheeler, T.J., Schreiber, F., Bateman, A., Eddy, S.R., 2015. HMMER
325 web server: 2015 update. *Nucleic Acids Res.* 43, W30–W38. doi:10.1093/nar/gkv397

326 Galperin, M.Y., Makarova, K.S., Wolf, Y.I., Koonin, E. V, 2015. Expanded microbial genome coverage and
327 improved protein family annotation in the COG database. *Nucleic Acids Res.* 43, D261-9.
328 doi:10.1093/nar/gku1223

329 Geiselbrecht, A.D., Hedlund, B.P., Tichi, M.A., Staley, J.T., 1998. Isolation of marine polycyclic aromatic
330 hydrocarbon (PAH)-degrading *Cycloclasticus* strains from the Gulf of Mexico and comparison of their PAH
331 degradation ability with that of puget sound *Cycloclasticus* strains. *Appl. Environ. Microbiol.* 64, 4703–10.

332 Golyshin, P.N., Werner, J., Chernikova, T.N., Tran, H., Ferrer, M., Yakimov, M.M., Teeling, H., Golyshina, O. V,
333 MAMBA Scientific Consortium, 2013. Genome Sequence of *Thalassolituus oleivorans* MIL-1 (DSM 14913T).
334 *Genome Announc.* 1, e00141113. doi:10.1128/genomeA.00141-13

335 Jeong, H., Yim, J.H., Lee, C., Choi, S.-H., Park, Y.K., Yoon, S.H., Hur, C.-G., Kang, H.-Y., Kim, D., Lee, H.H., Park, K.H.,
336 Park, S.-H., Park, H.-S., Lee, H.K., Oh, T.K., Kim, J.F., 2005. Genomic blueprint of *Hahella chejuensis*, a
337 marine microbe producing an algicidal agent. *Nucleic Acids Res.* 33, 7066–7073. doi:10.1093/nar/gki1016

338 Jukes T H, Cantor C R, 1969. Evolution of protein molecules, in: *Mammalian Protein Metabolism*, III. pp. 21–132.
339 doi:citeulike-article-id:768582

340 Kato, N., Yamazoe, K., Han, C.-G., Ohtsubo, E., 2003. New insertion sequence elements in the upstream region of
341 *cfiA* in imipenem-resistant *Bacteroides fragilis* strains. *Antimicrob. Agents Chemother.* 47, 979–85.

342 Kok, M., Oldenhuis, R., van der Linden, M.P., Raatjes, P., Kingma, J., van Lelyveld, P.H., Witholt, B., 1989. The
343 *Pseudomonas oleovorans* alkane hydroxylase gene. Sequence and expression. *J. Biol. Chem.* 264, 5435–41.

344 Konstantinidis, K.T., Tiedje, J.M., 2005. Towards a Genome-Based Taxonomy for Prokaryotes. *J. Bacteriol.* 187,
345 6258–6264. doi:10.1128/JB.187.18.6258-6264.2005

346 Kube, M., Chernikova, T.N., Al-Ramahi, Y., Beloqui, A., Lopez-Cortez, N., Guazzaroni, M.-E., Heipieper, H.J., Klages,
 347 S., Kotsyurbenko, O.R., Langer, I., Nechitaylo, T.Y., Lünsdorf, H., Fernández, M., Juárez, S., Ciordia, S.,
 348 Singer, A., Kagan, O., Egorova, O., Alain Petit, P., Stogios, P., Kim, Y., Tchigvintsev, A., Flick, R., Denaro, R.,
 349 Genovese, M., Albar, J.P., Reva, O.N., Martínez-Gomariz, M., Tran, H., Ferrer, M., Savchenko, A., Yakunin,
 350 A.F., Yakimov, M.M., Golyshina, O. V., Reinhardt, R., Golyshin, P.N., Petit, P.A., Stogios, P., Kim, Y.,
 351 Tchigvintsev, A., Flick, R., Denaro, R., Genovese, M., Albar, J.P., Reva, O.N., Martínez-Gomariz, M., Tran, H.,
 352 Ferrer, M., Savchenko, A., Yakunin, A.F., Yakimov, M.M., Golyshina, O. V., Reinhardt, R., Golyshin, P.N.,
 353 2013. Genome sequence and functional genomic analysis of the oil-degrading bacterium *Oleispira*
 354 *antarctica*. *Nat. Commun.* 4, 2156. doi:10.1038/ncomms3156

355 Lai, Q., Li, W., Wang, B., Yu, Z., Shao, Z., 2012. Complete genome sequence of the pyrene-degrading bacterium
 356 *Cycloclasticus* sp. strain P1. *J. Bacteriol.* 194, 6677. doi:10.1128/JB.01837-12

357 Lartigue, M.-F., Poirel, L., Aubert, D., Nordmann, P., 2006. In Vitro Analysis of ISEcp1B-Mediated Mobilization of
 358 Naturally Occurring β -Lactamase Gene blaCTX-M of *Kluyvera ascorbata*. *Antimicrob. Agents Chemother.* 50,
 359 1282–1286. doi:10.1128/AAC.50.4.1282-1286.2006

360 Lee, H.K., Chun, J., Moon, E.Y., Ko, S.H., Lee, D.S., Lee, H.S., Bae, K.S., 2001. *Hahella chejuensis* gen. nov., sp. nov.,
 361 an extracellular-polysaccharide-producing marine bacterium. *Int. J. Syst. Evol. Microbiol.* 51, 661–6.
 362 doi:10.1099/00207713-51-2-661

363 Lünsdorf, H., Abraham, W.-R., Yakimov, M.M., Timmis, K.N., Chernikova, T.N., Golyshin, P.N., 2002. *Oleiphilaceae*
 364 fam. nov., to include *Oleiphilus messinensis* gen. nov., sp. nov., a novel marine bacterium that obligately
 365 utilizes hydrocarbons. *Int. J. Syst. Evol. Microbiol.* 52, 901–911. doi:10.1099/00207713-52-3-901

366 McKew, B.A., Coulon, F., Osborn, A.M., Timmis, K.N., McGenity, T.J., 2007. Determining the identity and roles of
 367 oil-metabolizing marine bacteria from the Thames estuary, UK. *Environ. Microbiol.* 9, 165–76.
 368 doi:10.1111/j.1462-2920.2006.01125.x

369 Messina, E., Denaro, R., Crisafi, F., Smedile, F., Cappello, S., Genovese, M., Genovese, L., Giuliano, L., Russo, D.,
 370 Ferrer, M., Golyshin, P., Yakimov, M.M., 2016. Genome sequence of obligate marine polycyclic aromatic
 371 hydrocarbons-degrading bacterium *Cycloclasticus* sp. 78-ME, isolated from petroleum deposits of the

372 sunken tanker Amoco Milford Haven, Mediterranean Sea. *Mar. Genomics* 25, 11–13.
373 doi:10.1016/j.margen.2015.10.006

374 Murdoch, S.L., Trunk, K., English, G., Fritsch, M.J., Pourkarimi, E., Coulthurst, S.J., 2011. The opportunistic
375 pathogen *Serratia marcescens* utilizes type VI secretion to target bacterial competitors. *J. Bacteriol.* 193,
376 6057–69. doi:10.1128/JB.05671-11

377 Panke, S., Röthlisberger, M., van Beilen, J.B., Lucchini, S., Witholt, B., Franchini, A.G., 2001. Analysis of
378 *Pseudomonas putida* alkane-degradation gene clusters and flanking insertion sequences: evolution and
379 regulation of the alk genes. *Microbiology* 147, 1621–1630. doi:10.1099/00221287-147-6-1621

380 Sabirova, J.S., Becker, A., Lünsdorf, H., Nicaud, J.-M., Timmis, K.N., Golyshin, P.N., 2011. Transcriptional profiling
381 of the marine oil-degrading bacterium *Alcanivorax borkumensis* during growth on n-alkanes. *FEMS*
382 Microbiol. Lett. 319, 160–8. doi:10.1111/j.1574-6968.2011.02279.x

383 Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees.
384 *Mol. Biol. Evol.* 4, 406–25.

385 Schneiker, S., Martins dos Santos, V.A.P., Bartels, D., Bekel, T., Brecht, M., Buhrmester, J., Chernikova, T.N.,
386 Denaro, R., Ferrer, M., Gertler, C., Goesmann, A., Golyshina, O. V, Kaminski, F., Khachane, A.N., Lang, S.,
387 Linke, B., McHardy, A.C., Meyer, F., Nechitaylo, T., Pühler, A., Regenhardt, D., Rupp, O., Sabirova, J.S.,
388 Selbitschka, W., Yakimov, M.M., Timmis, K.N., Vorhölter, F.-J., Weidner, S., Kaiser, O., Golyshin, P.N., 2006.
389 Genome sequence of the ubiquitous hydrocarbon-degrading marine bacterium *Alcanivorax borkumensis*.
390 *Nat. Biotechnol.* 24, 997–1004. doi:10.1038/nbt1232

391 Sievers, F., Higgins, D.G., 2014. Clustal Omega, in: *Current Protocols in Bioinformatics*. John Wiley & Sons, Inc.,
392 Hoboken, NJ, USA, p. 3.13.1-3.13.16. doi:10.1002/0471250953.bi0313s48

393 Soki, J., Gal, M., Brazier, J.S., Rotimi, V.O., Urbán, E., Nagy, E., Duerden, B.I., 2006. Molecular investigation of
394 genetic elements contributing to metronidazole resistance in *Bacteroides* strains. *J. Antimicrob.*
395 *Chemother.* 57, 212–220. doi:10.1093/jac/dki443

396 Sone, Y., Nakamura, R., Pan-Hou, H., Itoh, T., Kiyono, M., 2013. Role of MerC, MerE, MerF, MerT, and/or MerP in
397 resistance to mercurials and the transport of mercurials in *Escherichia coli*. *Biol. Pharm. Bull.* 36, 1835–41.

398 Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies.
399 *Bioinformatics* 30, 1312–3. doi:10.1093/bioinformatics/btu033

400 Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: Molecular Evolutionary Genetics
401 Analysis Version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi:10.1093/molbev/mst197

402 Throne-Holst, M., Wentzel, A., Ellingsen, T.E., Kotlar, H.-K., Zotchev, S.B., 2007. Identification of Novel Genes
403 Involved in Long-Chain n-Alkane Degradation by *Acinetobacter* sp. Strain DSM 17874. *Appl. Environ.*
404 *Microbiol.* 73, 3327–3332. doi:10.1128/AEM.00064-07

405 Treves, D.S., Manning, S., Adams, J., 1998. Repeated evolution of an acetate-crossfeeding polymorphism in long-
406 term populations of *Escherichia coli*. *Mol. Biol. Evol.* 15, 789–97.

407 Vandecraen, J., Chandler, M., Aertsen, A., Van Houdt, R., 2017. The impact of insertion sequences on bacterial
408 genome plasticity and adaptability. *Crit. Rev. Microbiol.* 1–22. doi:10.1080/1040841X.2017.1303661

409 Varani, A.M., Siguier, P., Goubeyre, E., Charneau, V., Chandler, M., 2011. ISsaga is an ensemble of web-based
410 methods for high throughput identification and semi-automatic annotation of insertion sequences in
411 prokaryotic genomes. *Genome Biol.* 12, R30. doi:10.1186/gb-2011-12-3-r30

412 Wentzel, A., Ellingsen, T.E., Kotlar, H.-K., Zotchev, S.B., Throne-Holst, M., 2007. Bacterial metabolism of long-
413 chain n-alkanes. *Appl. Microbiol. Biotechnol.* 76, 1209–1221. doi:10.1007/s00253-007-1119-1

414 Wuichet, K., Alexander, R.P., Zhulin, I.B., 2007. Comparative Genomic and Protein Sequence Analyses of a
415 Complex System Controlling Bacterial Chemotaxis, in: *Methods in Enzymology*. pp. 3–31.
416 doi:10.1016/S0076-6879(06)22001-9

417 Yakimov, M.M., Golyshin, P.N., Lang, S., Moore, E.R., Abraham, W.R., Lünsdorf, H., Timmis, K.N., 1998.
418 *Alcanivorax borkumensis* gen. nov., sp. nov., a new, hydrocarbon-degrading and surfactant-producing
419 marine bacterium. *Int. J. Syst. Bacteriol.* 48 Pt 2, 339–48. doi:10.1099/00207713-48-2-339

420 Yakimov, M.M., Giuliano, L., Denaro, R., Crisafi, E., Chernikova, T.N., Abraham, W.-R., Luensdorf, H., Timmis, K.N.,
 421 Golyshin, P.N., 2004. *Thalassolituus oleivorans* gen. nov., sp. nov., a novel marine bacterium that obligately
 422 utilizes hydrocarbons. *Int. J. Syst. Evol. Microbiol.* 54, 141–8. doi:10.1099/ijs.0.02424-0

423 Yakimov, M.M., Giuliano, L., Gentile, G., Crisafi, E., Chernikova, T.N., Abraham, W.-R., Lünsdorf, H., Timmis, K.N.,
 424 Golyshin, P.N., 2003. *Oleispira antarctica* gen. nov., sp. nov., a novel hydrocarbonoclastic marine bacterium
 425 isolated from Antarctic coastal sea water. *Int. J. Syst. Evol. Microbiol.* 53, 779–85. doi:10.1099/ijs.0.02366-0

426 Yakimov, M.M., Timmis, K.N., Golyshin, P.N., 2007. Obligate oil-degrading marine bacteria. *Curr. Opin.*
 427 *Biotechnol.* 18, 257–266. doi:10.1016/j.copbio.2007.04.006

428

429

430